一般のコスト関数に対する ハフマン木問題及び探索木問題

藤原 洋志 (豊橋技術科学大学)
Tobias Jacobs (国立情報学研究所)

ハフマン木問題

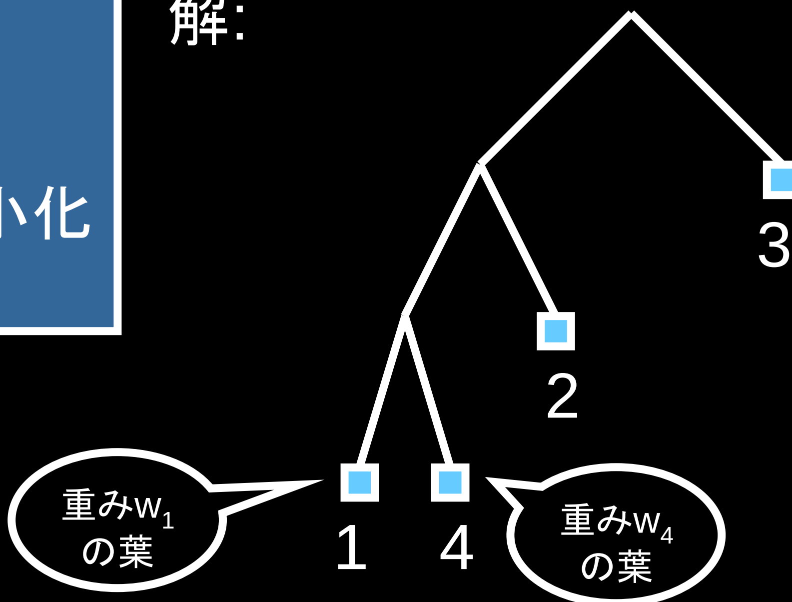
入力: 重み列 $w_1, w_2, ..., w_n$
出力: 2分木
目的:

$$\sum_{i=1}^{n} w_i d_i \rightarrow \text{最小化}$$

d_i: i番目の葉の深さ

例: $w_1 = 1, w_2 = 3, w_3 = 5, w_4 = 1$

解:



目的関数値 = 3 + 3 + 6 + 5 = 17

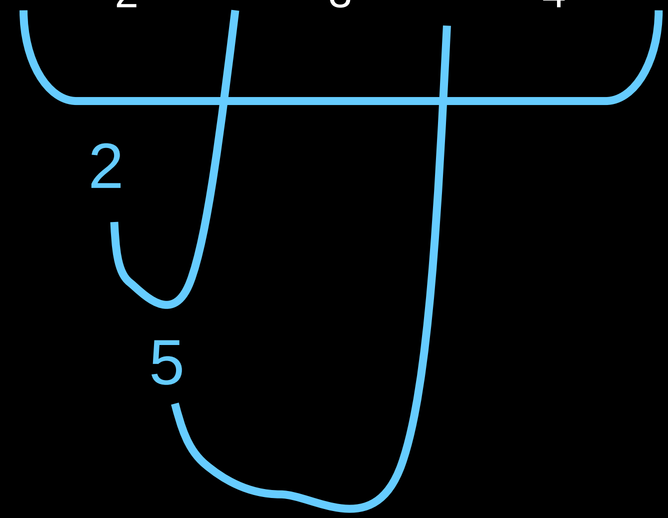
ハフマン木問題

入力: 重み列 $w_1, w_2, ..., w_n$
出力: 2分木
目的:

$$\sum_{i=1}^{n} w_i d_i \rightarrow \text{最小化}$$

d_i: i番目の葉の深さ

例: $w_1 = 1, w_2 = 3, w_3 = 5, w_4 = 1$



Huffmanアルゴリズム

ハフマン木問題

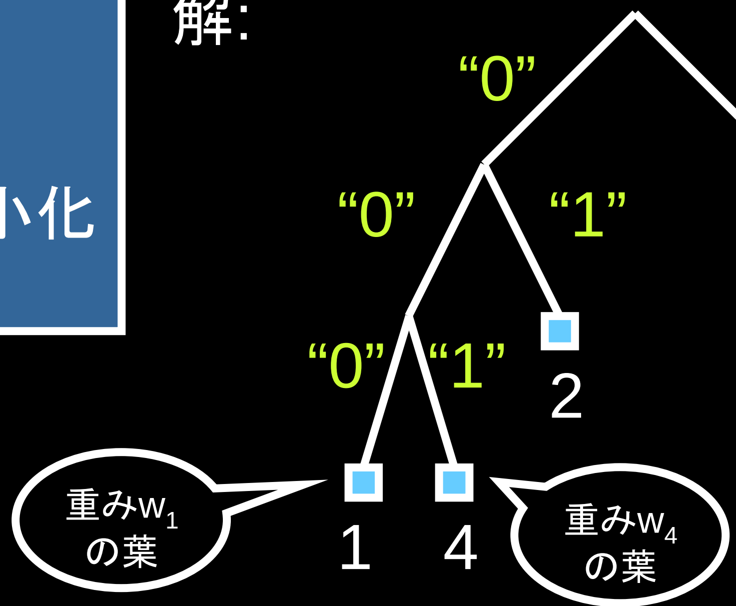
入力: 重み列 $w_1, w_2, \ldots, w_n$
出力: 2分木
目的:

$$\sum_{i=1}^{n} w_i d_i \rightarrow \text{最小化}$$

d_i: i番目の葉の深さ

例: $w_1 = 1, w_2 = 3, w_3 = 5, w_4 = 1$

解:



ハフマン符号:
1: "000"
2: "01"
3: "1"
4: "001"

探索木問題

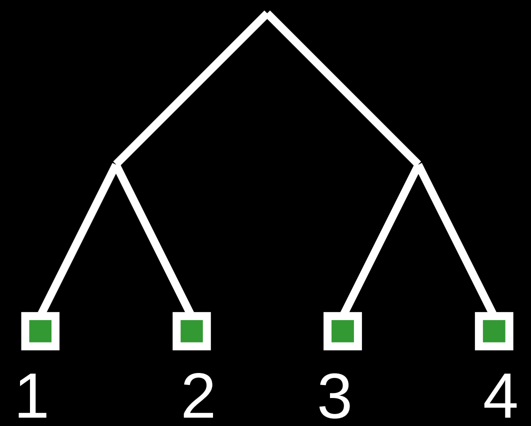
入力: 重み列 $w_1, w_2, ..., w_n$
出力: 2分木 (葉順そのまま)
目的:

$$\sum_{i=1}^{n} w_i d_i \rightarrow \text{最小化}$$

d_i: i番目の葉の深さ

例: $w_1 = 1, w_2 = 3, w_3 = 5, w_4 = 1$

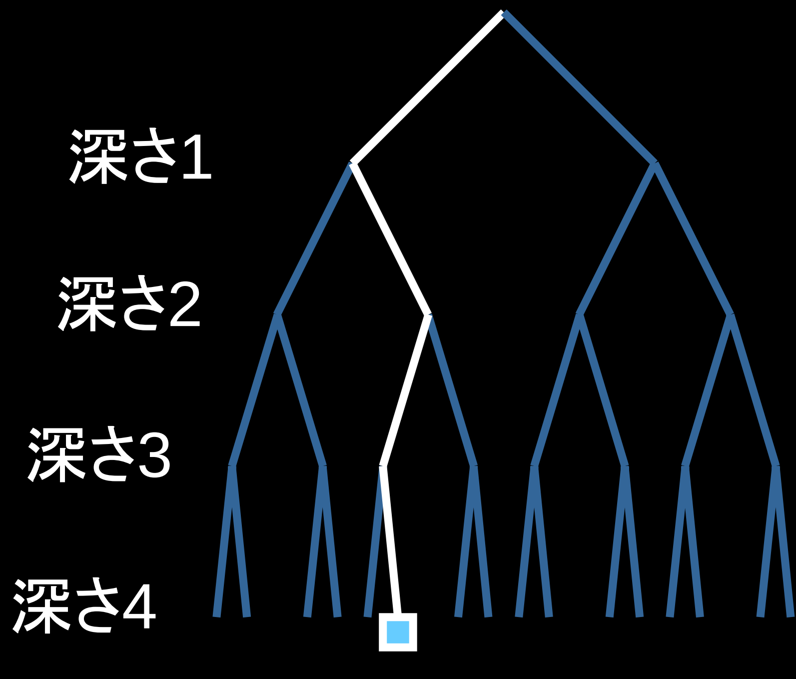
解:



$w_1, w_2, ..., w_n$の順に葉が並ぶ

目的関数値 = 2 + 6 + 10 + 2 = 20

コスト関数の一般化



目的関数: $\sum_{i=1}^{n} w_i d_i$

1つの葉の寄与分 $w_i d_i$
これは<u>深さに比例</u>

一般化

目的関数: $\sum_{i=1}^{n} f_i(d_i)$

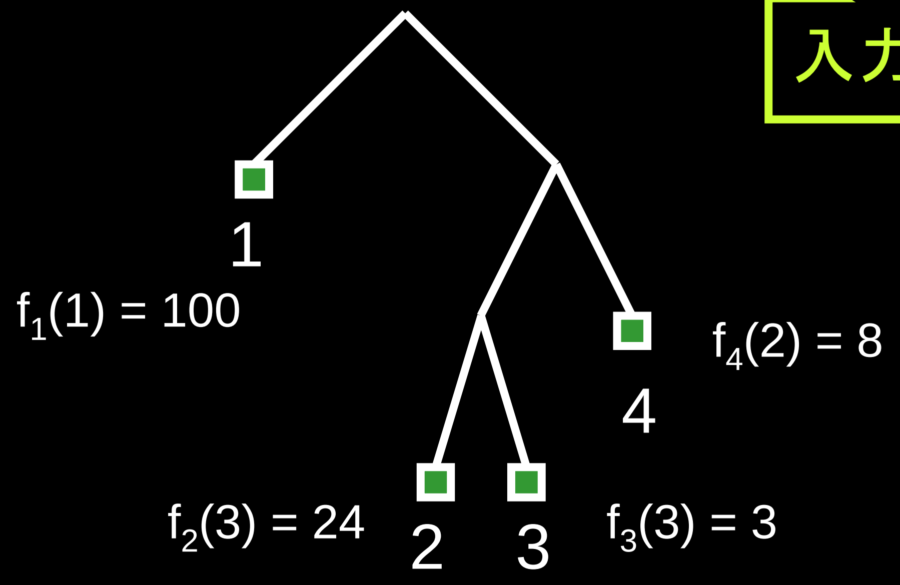
1つの葉の寄与分 $f_i(d_i)$

一般のコスト関数に対する探索木問題

例: $f_1(x) = 100x,\ f_2(x) = 8x,\ f_3(x) = x,\ f_4(x) = x^3$

入力は関数列

解:



目的関数値 = 100 + 24 + 3 + 8 = 135

ハフマン木問題

入力: 重み列 $w_1, w_2, \ldots, w_n$
出力: 2分木
目的:
$$\sum_{i=1}^{n} w_i d_i$$

一般化

$$\max_{1\le i\le n} f_i(d_i)$$ もOK

一般のコスト関数に対するハフマン木問題

入力: 関数列 $f_1, f_2, \ldots, f_n$
出力: 2分木
目的:
$$\sum_{i=1}^{n} f_i(d_i) \rightarrow \text{最小化}$$

探索木問題

入力: 重み列 $w_1, w_2, \ldots, w_n$
出力: 2分木 (葉順そのまま)
目的:
$$\sum_{i=1}^{n} w_i d_i$$

一般化

$$\max_{1\le i\le n} f_i(d_i)$$ もOK

一般のコスト関数に対する探索木問題

入力: 関数列 $f_1, f_2, \ldots, f_n$
出力: 2分木 (葉順そのまま)
目的:
$$\sum_{i=1}^{n} f_i(d_i) \rightarrow \text{最小化}$$

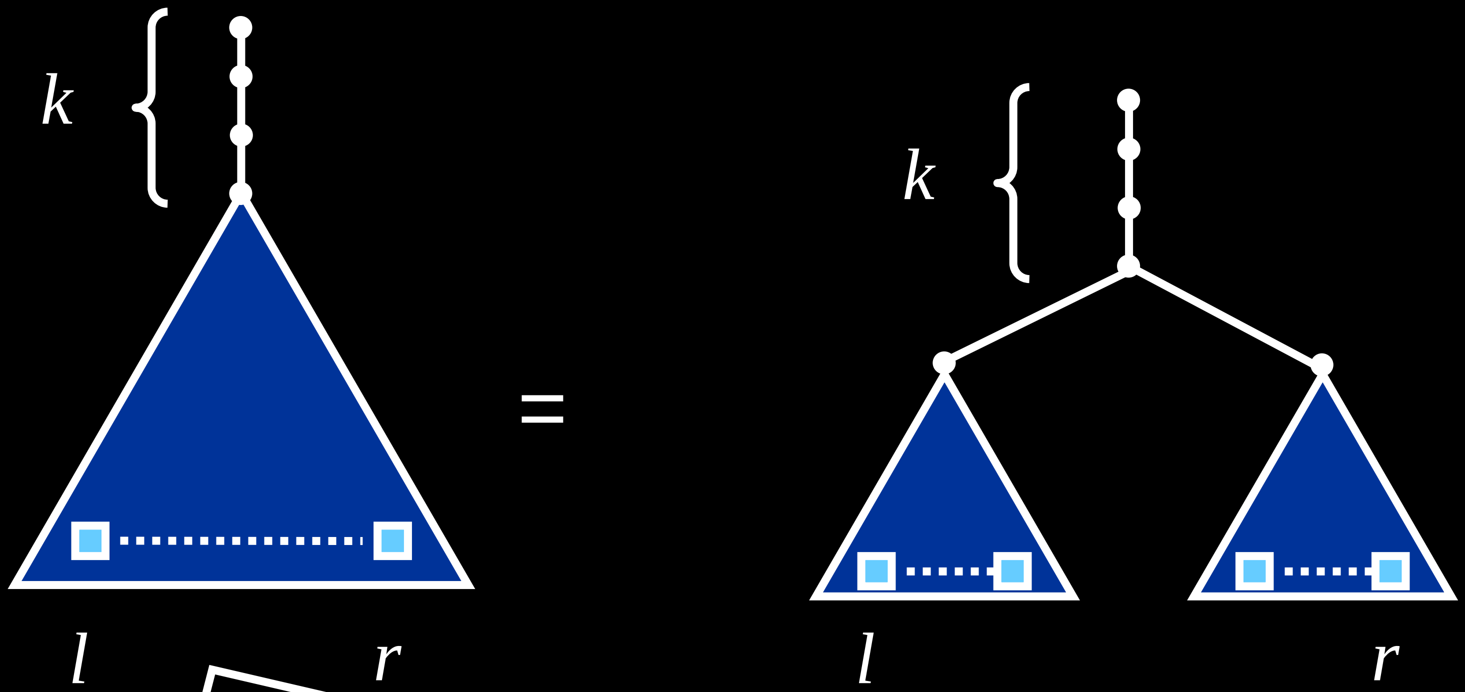
既存研究

- ハフマン木問題 $f_i(x) = w_i\, x$
  - Huffmanアルゴリズム O(n log n) [Huffman1952]
- 探索木問題 $f_i(x) = w_i\, x$
  - 動的計画法 $O(n^3)$ [Gilbert&Moore1959]
  - 動的計画法 $O(n^2)$ [Knuth1971]
  - Hu-Tuckerアルゴリズム O(n log n) [Hu&Tucker1971]
- 探索木問題 $f_i(x) = w_i\, a^x$
  - $a>1$の場合に対してHu-Tuckerアルゴリズム O(n log n) [Hu&Tucker&Tamaki1979]
  - $0<a<1$の場合に対して動的計画法 $O(n^3)$ [Baer2010]

本研究の成果概要

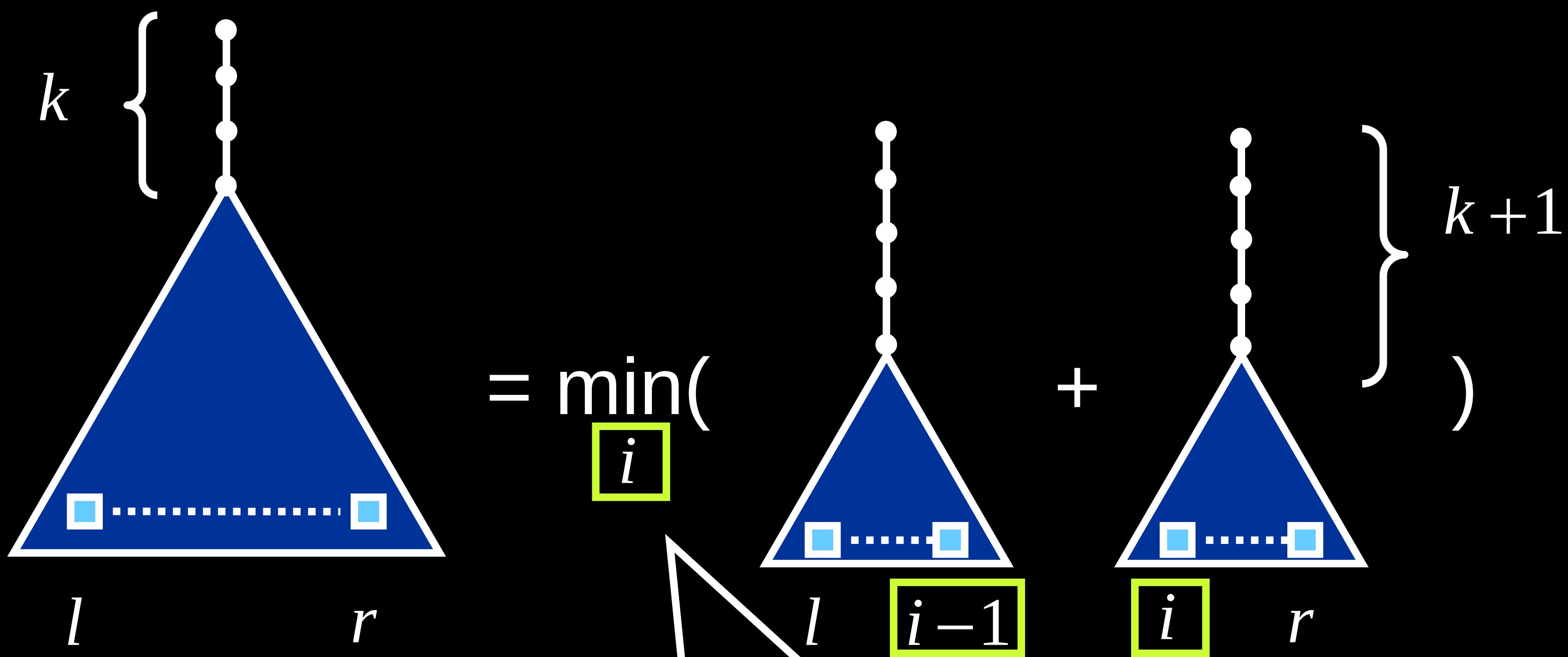
- 探索木問題に対する動的計画法
  - 任意の$\{f_i\}$ $O(n^4)$
  - $\{f_i\}$が全て非減少かつ凸 ⇒ 構造的連続性 $O(n^3)$
  - 目的関数がmax f_i で$\{f_i\}$が全て非減少 ⇒ 構造的連続性 $O(n^3)$
- ハフマン木問題
  - $P \neq NP$ ⇒ $\forall \alpha(n)$-近似困難
  - 目的関数がmax f_i で$\{f_i\}$が全て単調⇒$O(n^3 \log n)$

探索木問題に対する動的計画法



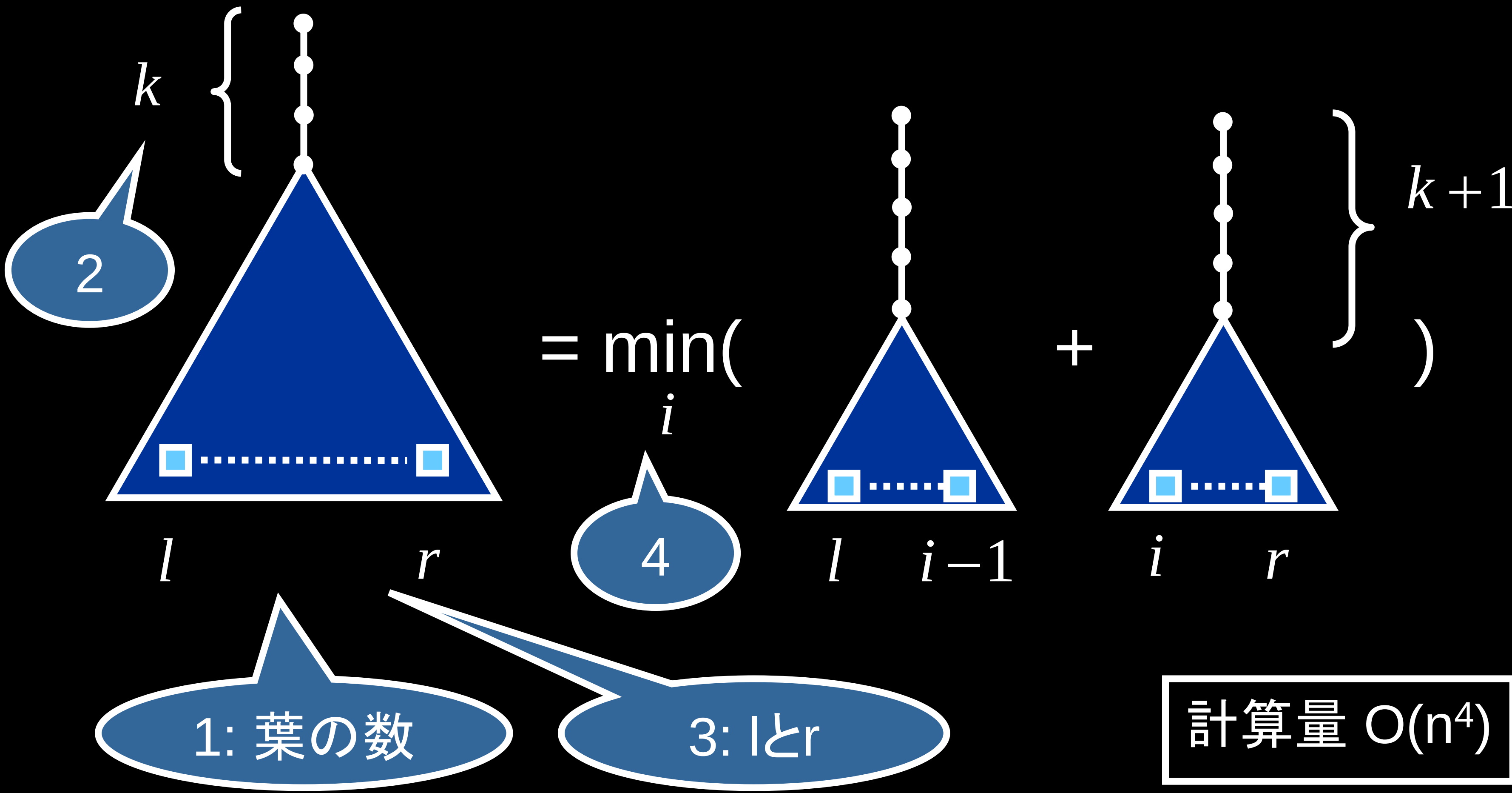
部分問題(l,r,k)(つまり、lからr番目までの葉+長さkのパス)に対する最適解

探索木問題に対する動的計画法



コストが最小となるものを
選ぶ

探索木問題に対する動的計画法



計算量 $O(n^4)$

探索木問題に対する動的計画法の計算量

任意の$\{f_i\}$

$O(n^4)$

構造的連続性

$O(n^3)$

定理

$\{f_i\}$が全て凸かつ非減少
⇒
構造的連続性

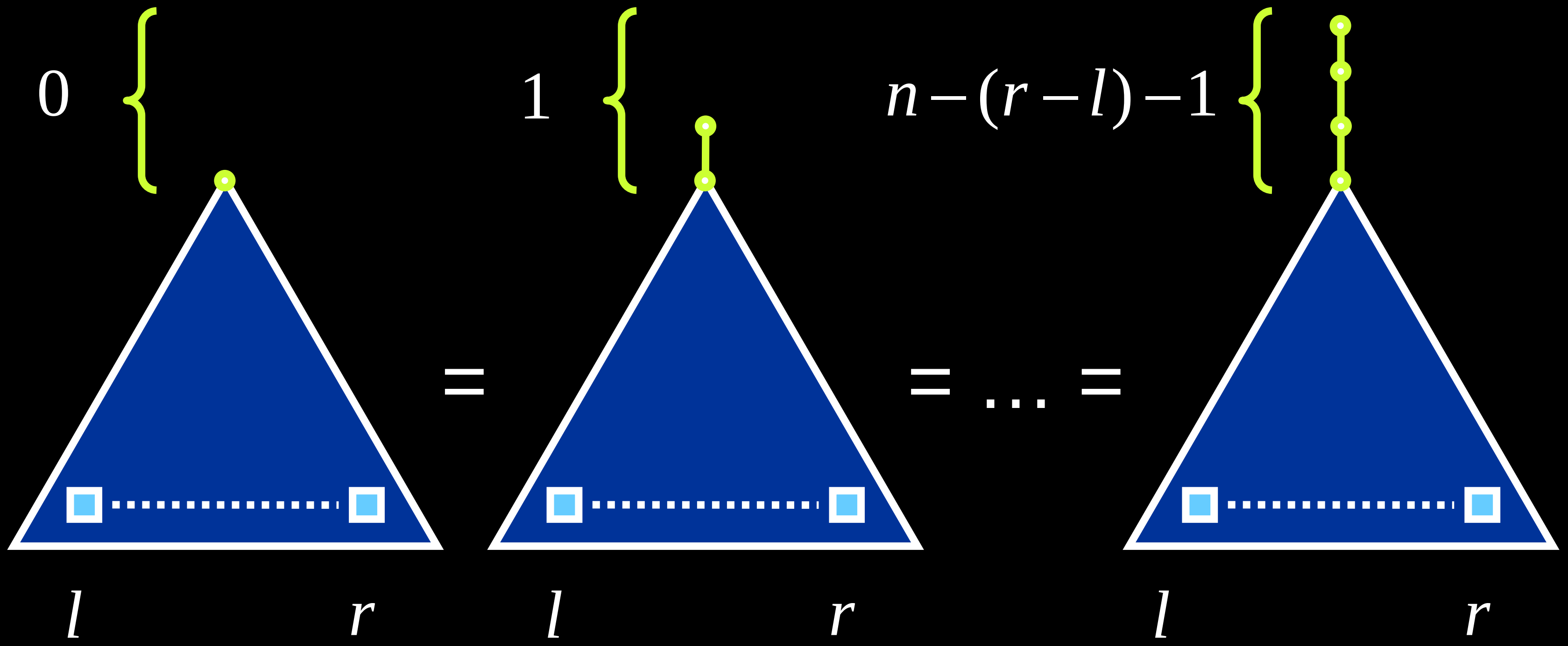
部分木最適性

$O(n^2)$

$f_i(x) = w_i x$ や $f_i(x) = w_i a^x$ (a>1)に対して
Hu-Tuckerアルゴリズム
$O(n \log n)$

$O(n^3)$

部分木最適性とは



パスの長さに関係なく同じ部分木が最適
異なるkについて再計算不要

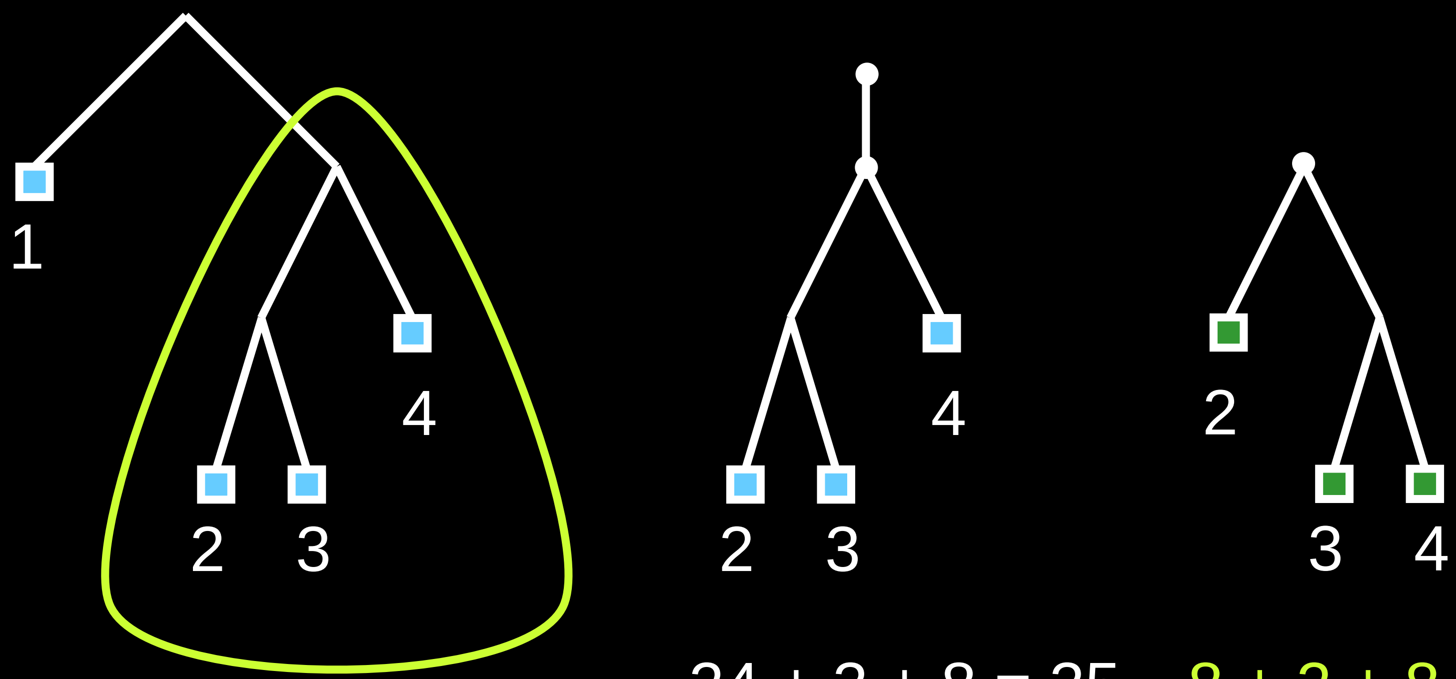
計算量 O(n^3)

部分木最適性を満たす$\{f_i\}$

- $f_i(x) = w_i x$
- $f_i(x) = w_i a^x$

部分木最適性を満たさない$\{f_i\}$

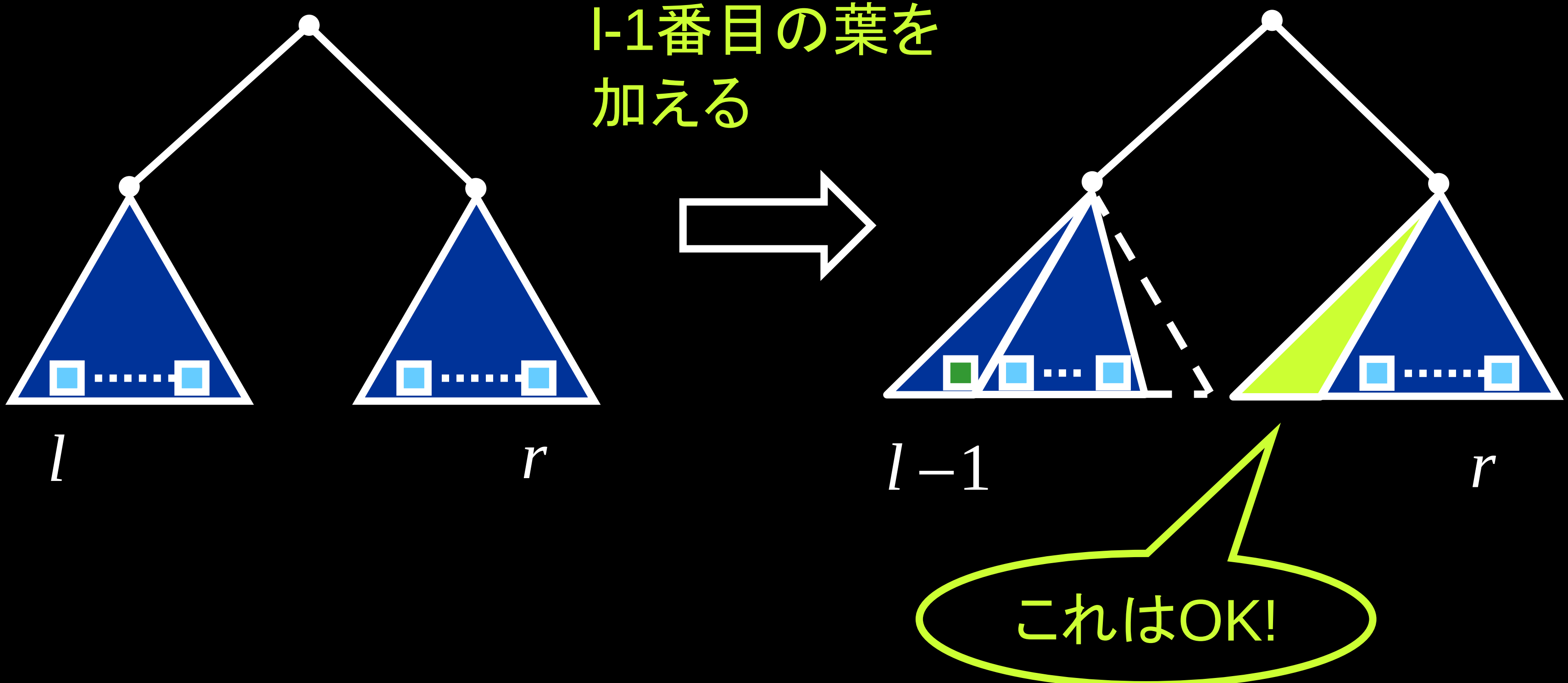
例: $f_1(x) = 100x,\ f_2(x) = 8x,\ f_3(x) = x,\ f_4(x) = x^3$



24 + 3 + 8 = 35　8 + 2 + 8 = 18

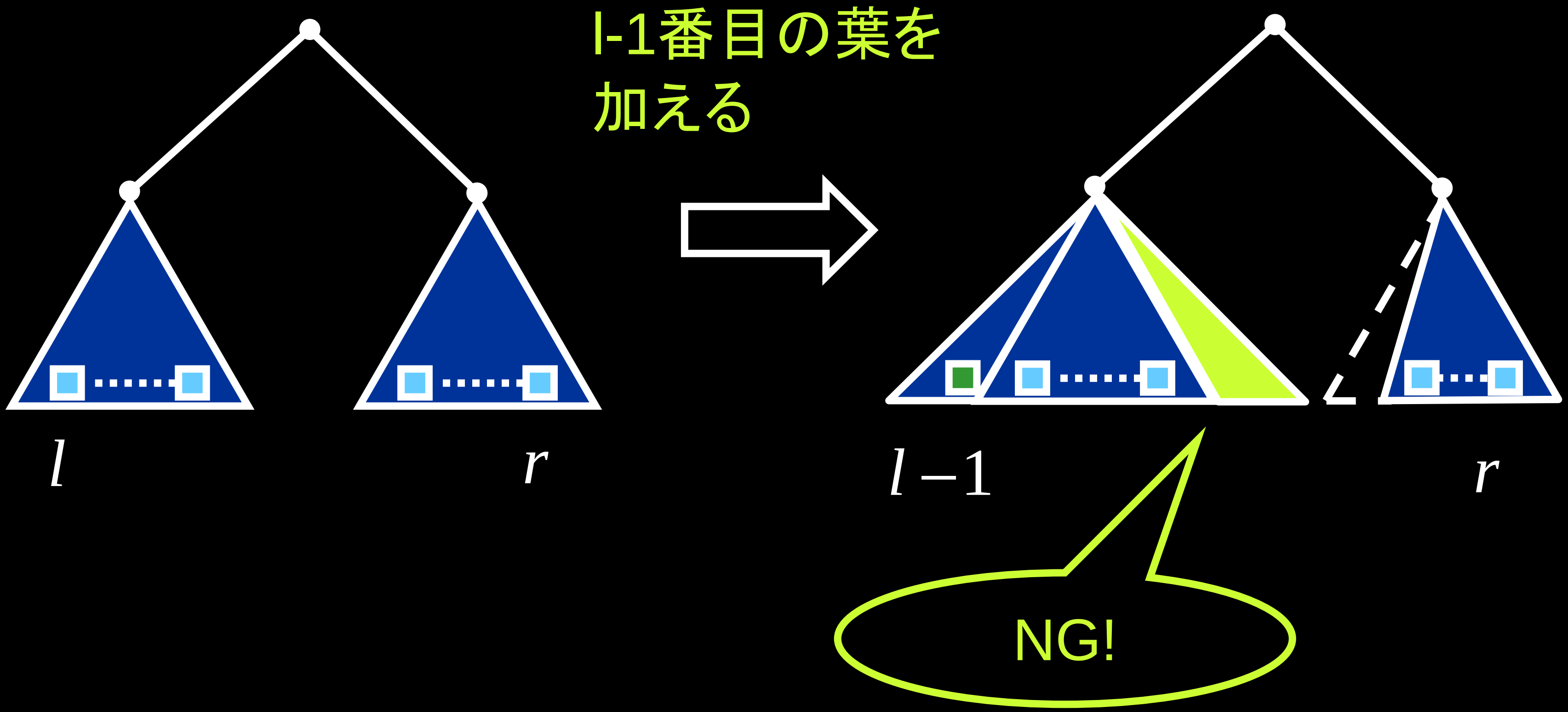
構造的連続性とは: 大雑把な説明

部分問題(l,r,k)に対する最適解　　　　部分問題(**l-1**,r,k)に対する最適解

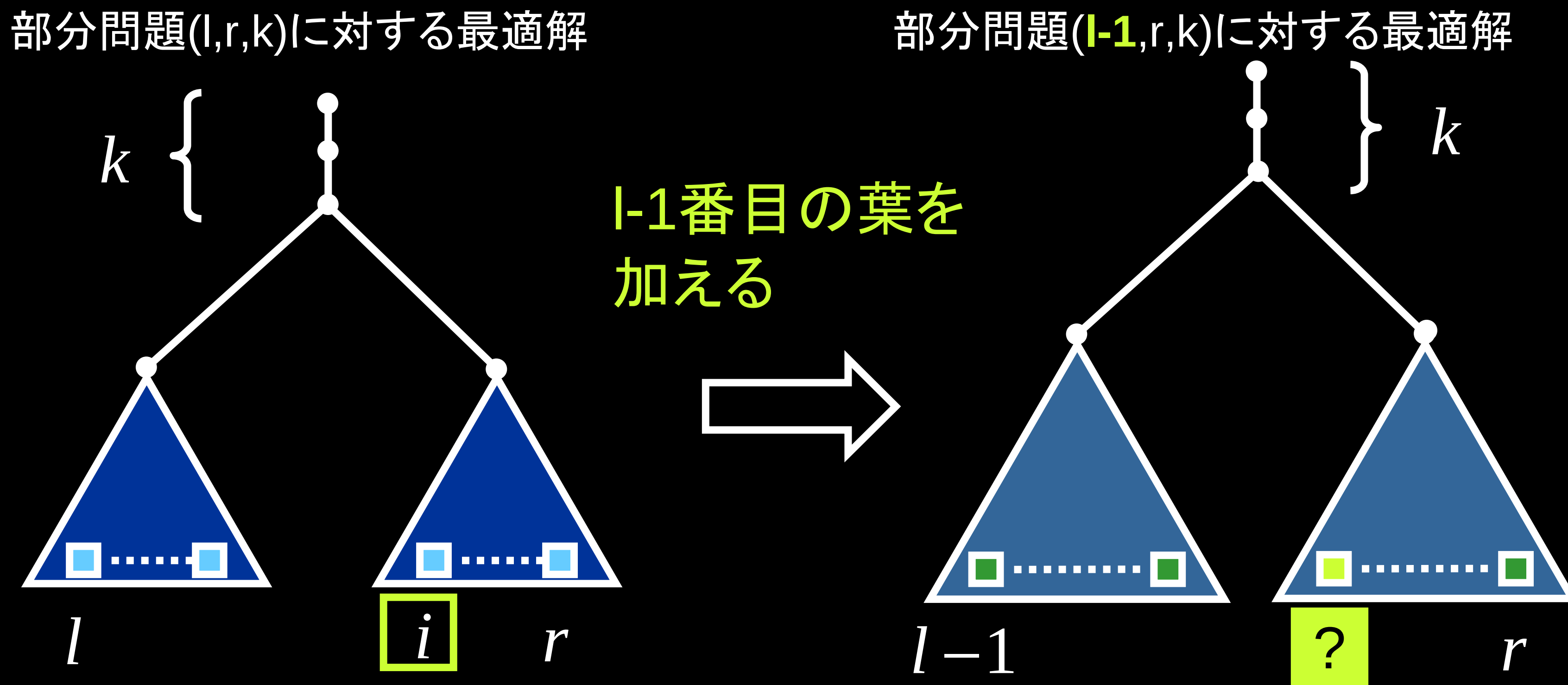


構造的連続性とは:
大雑把な説明

部分問題(l,r,k)に対する最適解　　　　部分問題(l-1,r,k)に対する最適解



構造的連続性とは

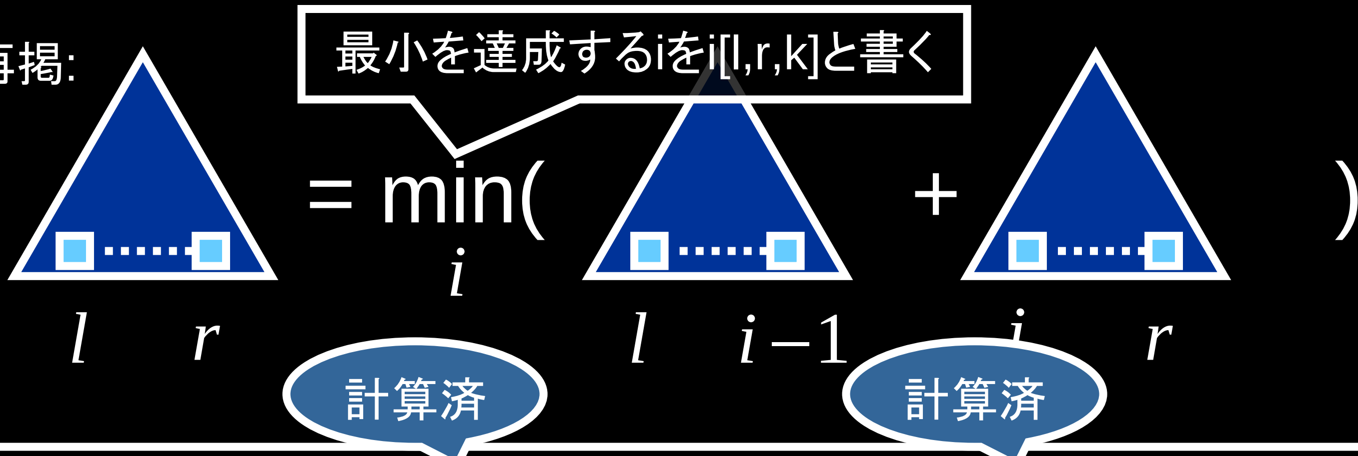


定義

構造的連続性とは、全ての部分問題に対して、この葉がi番目以下となるような最適な部分木が存在 (かつ左右対称の命題が真)

構造的連続性がどう役に立つのか？

再帰関係再掲:



構造的連続性 ⇒ i[l,r-1,k] ≦ i[l,r,k] ≦ i[l+1,r,k]

部分問題(1,a,k),(2,a+1,k),…,(n-a+1,n,k)に対しiの繰り返し回数は

$$\sum_{l=1}^{n-a}(i[l+1,l+a,k]-i[l,l+a-1,k]+1)$$

$$= i[n-a+1,n,k]-i[1,a,k]+n-a$$

$$\leq 2n$$

計算量 O(n^3)

構造的連続性を満たす$\{f_i\}$

- $f_i(x) = w_i\, x$
- 我々は次の十分条件を与える

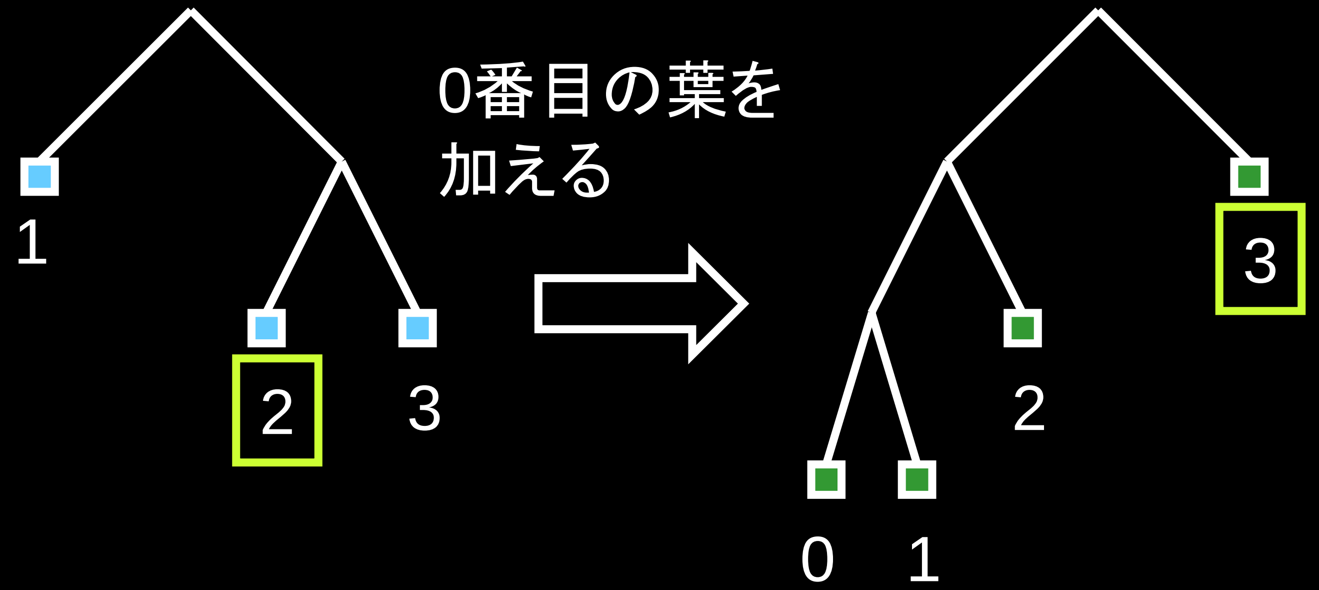
定理

$\{f_i\}$が全て非減少かつ凸 ⇒ 構造的連続性

証明は省略

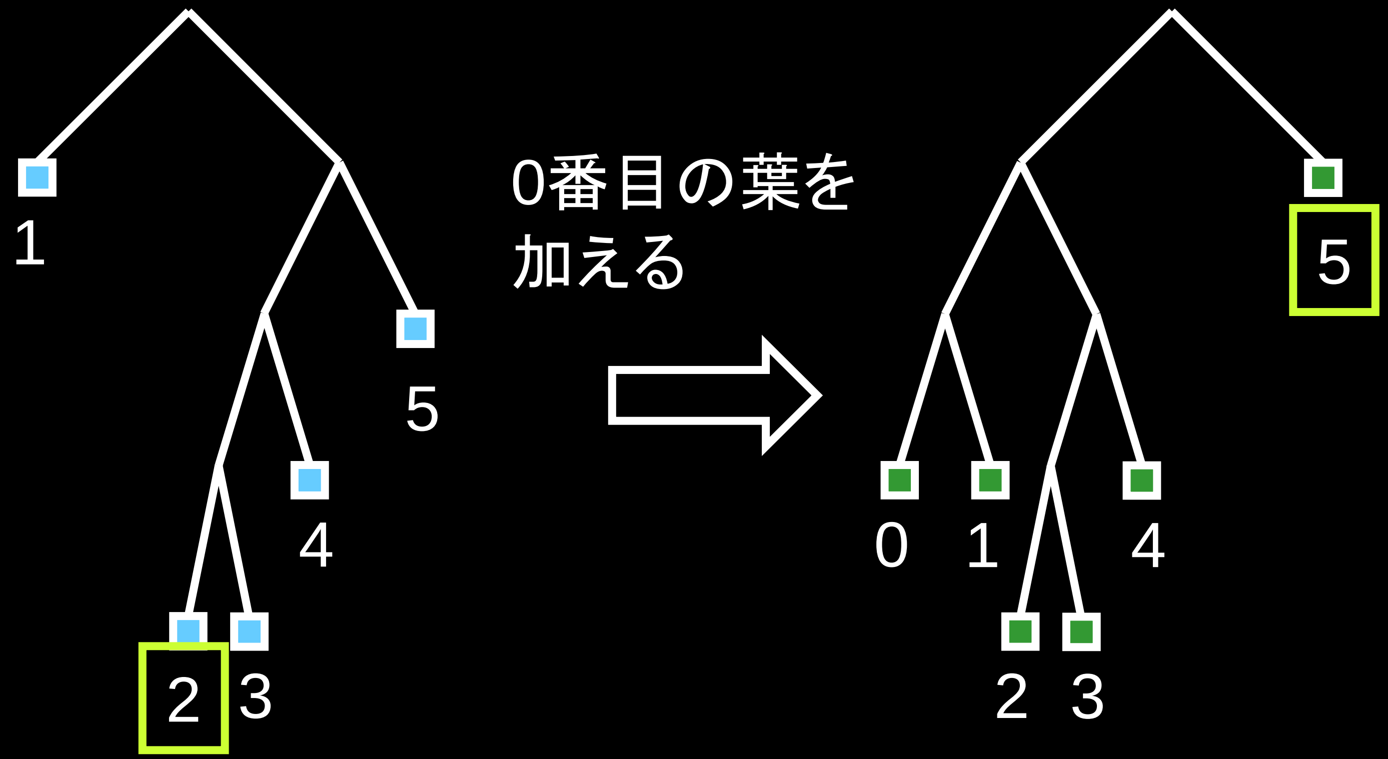
部分木最適性を満たすが構造的連続性を満たさない$\{f_i\}$

例: $f_0(x) = 0$, $f_1(x) = 1.5 \cdot (1-2^{-x})$,
$f_2(x) = (1-2^{-x})$, $f_3(x) = (1-2^{-x})$



構造的連続性も部分木最適性も満たさない$\{f_i\}$

例: $f_0(x) = 0$, $f_1(x) = 1500(1\text{-}2^{-x})$, $f_2(x) = 8x$, $f_3(x) = x$,
$f_4(x) = x^3$, $f_5(x) = 1000(1\text{-}2^{-x})$



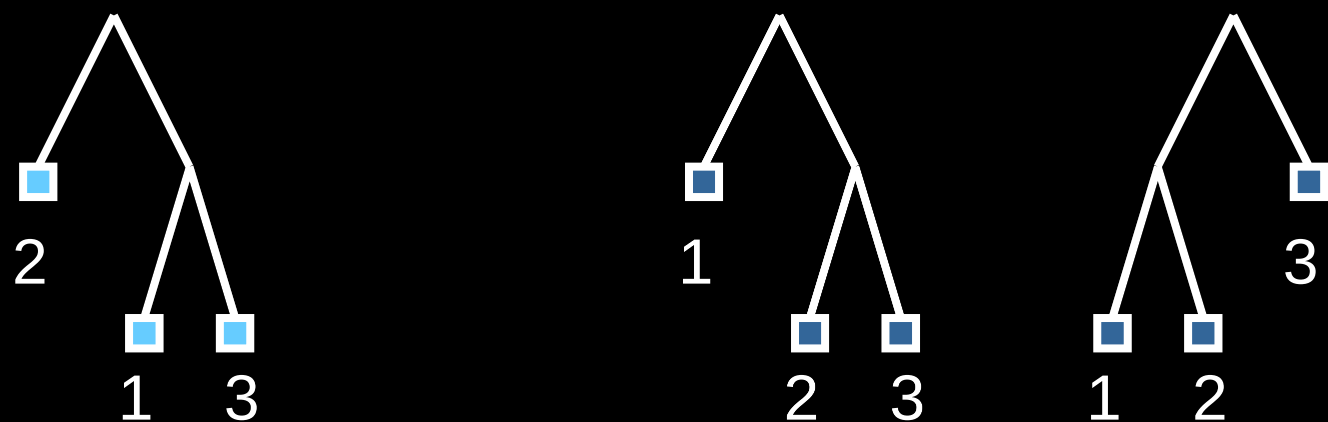
ハフマン木問題

- $f_i(x) = w_i x$ に対しては「ソート + 探索木問題のアルゴリズム」により正しい解
- 一般には不成立

値域に順序

例: $f_1(x) = x-1$, $f_2(x) = 8x-6$, $f_3(x) = x+10$

解:



2 + 1 + 11 = 14　　0+10+11=22　　1+10+10=20

ハフマン木問題の複雑さ

- 探索木問題と比べて格段に難しい

定理

$P \neq NP \Rightarrow \forall \alpha(n)$-近似困難

証明は省略: Exact Cover by 3-Setsを帰着

- ただし、目的関数がmax f_i で$\{f_i\}$が全て単調 $\Rightarrow O(n^3 \log n)$

まとめ

- 探索木問題に対する動的計画法
  - 任意の$\{f_i\}$　$O(n^4)$
  - $\{f_i\}$が全て非減少かつ凸 ⇒ 構造的連続性 ⇒ $O(n^3)$
- ハフマン木問題
  - $P \neq NP$ ⇒ $\forall \alpha(n)$-近似困難
  - 目的関数が$\max f_i$で$\{f_i\}$が全て単調 ⇒ $O(n^3 \log n)$

今後の課題

- 探索木問題
  - 計算量下界
  - k分木
- ハフマン木
  - 多項式時間で解けるクラス